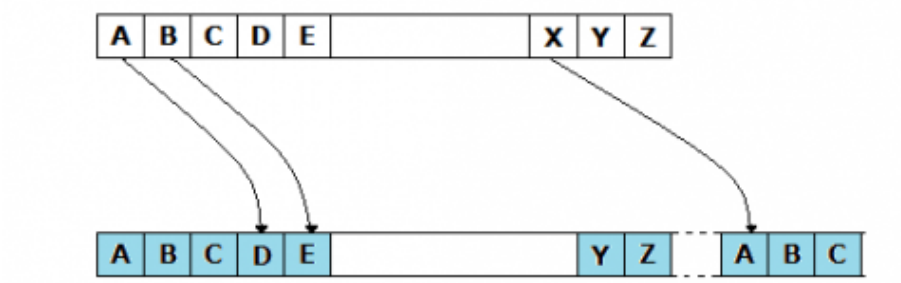
Prime Numbers in Cryptography

The reason all the money in your bank account is safe is because of… prime numbers? Those numbers you learnt about in primary school that are divisible by only 1 and themselves, are keeping me safe? The number 7 is keeping the 7 digits in my bank account secure? To some extent, yes, they are, although perhaps not quite as simply as the prime number 7 alone protecting millions of pounds. Although, the principles are the same, and I have taken on this project to explore how a concept which was discovered millennia ago- which may have been seen as almost useless- has fascinatingly become an essential part of our everyday life thousands of years later.

Although cryptography may appear like a fairly new science, seemingly necessary only to aid and protect the gain in popularity of widespread use of machines and computers, the concept has actually been around for quite some time, though presumably not in the form in which we understand today, and certainly not using prime numbers.

While perhaps a stretch, the first use of cryptography was in 1900BC Egypt, where the tomb of Khnumhotep II used unordinary hieroglyphics. However, it is thought that the purpose of this was to make the text appear more dignified and sophisticated rather than to conceal the meaning of a message, as conventionally believed. Nonetheless, it is still regarded as a transformation of text, and so is considered the first instance of encryption. Encryption is the process of scrambling data to make it illegible to unauthorised viewers

More iconically, however, was Julius Caesar's Caesar Cipher. Much like we will see later on, Caesar needed to communicate with his army generals on the frontlines, which allowed his empire to flourish as it did. His solution to do so, while remaining undetectable to unwanted eyes, would be to use a substitution cipher, which would 'shift' the alphabet by a certain amount to encode a message. The characters would also wrap around at the end, so 'X' would be replaced by 'A'. It would look as such:



The word 'base', with a key of 3 to the right, would appear as 'edvh'. However, the concept of keys was not properly introduced until the 14th century. This would mean that anyone that saw this message would deem it nonsensical, but with an idea that the alphabet had been shifted, the message can be toyed with and eventually deciphered (decrypted), allowing Caesar's troops to act on his command. This does, however, not completely address the problem of interception. Should those unwanted eyes also catch on that the letters were simply substituted, decrypting the message first would be a matter of luck, as neither party had access to a key.

So to keep up with the times, along with the rise of electronics, Hebern made a contraption at the dawn of the 19th century: the Hebern rotor machine. It used a single rotor, in which the secret key is embedded in a rotating disc. The key encoded a substitution table and each key press from the keyboard resulted in the output of cipher text. This also rotated the disc by one notch and a different table would then be used for the next plain text character. The later used Engima machine was invented by the Germans at the end of World War I, and was heavily used by the German forces during the Second World War, to communicate with their forces and generals. The Enigma machine used 3 or 4 or even more rotors. The rotors rotate at different rates as you type on the keyboard and output appropriate letters of cipher text. In this case the key was the initial setting of the rotors. This was inevitably decrypted, proving it was not a suitable method of encryption in the long term.

Up to the Second World War, most of the work on cryptography was for military purposes, usually used to hide secret military information. However, cryptography attracted commercial attention post-war, with businesses trying to secure their data from competitors, and, in turn, communicate discreetly. A necessity for encryption arose.

So, how do prime numbers actually come in to all of this. Well, we first must have an understanding of infinitely many prime numbers existing. Conveniently, however, the Greeks showed us this about 2,000 years ago.

First, we must make an assumption that there are a finite amount of prime numbers, which can be set out in a list as such:

p1, p2, p3, … pn-1, pn

Then, we can consider what will happen if we let some number N, be the product of every item in the list of the prime numbers, and add 1 to said product.

N= p1 x p2 x p3 x pn-1 x pn (+1)

We can now consider 2 cases: N either is or is not a prime number

If N is a prime number, that means that our original assumption is incorrect, and there are in fact more prime numbers than given in the list.

If N is not a prime number, that must mean that it is divisible by some other prime number. And, by the nature in which we constructed this number, we know that it cannot be made up of any of the primes in the original list. Additionally, given that every non-prime number can be rewritten using prime factorisation, the only numbers that could make up N are primes that are larger than the greatest primes in the original list, as the ones in the list cannot be factors, due to the '+1". Therefore, this once again disproves the original statement, as we have discovered that there must be a prime number greater than pn.

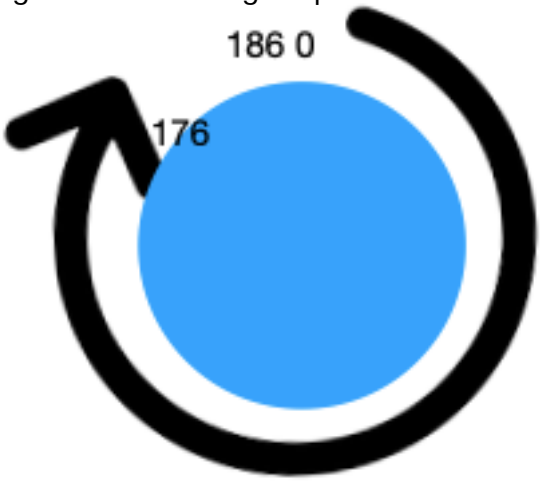
This type of proof is called a proof by contradiction, or as the Greeks called it, r*eductio ad absurdum*, meaning proof by absurdity. Although this likely had little use to the Greeks who had discovered it, other than to, of course, entertain the ambitious minds of mathematicians of the time, this fact serves us very importantly today.

Diverting our attention back to prime numbers, there are two fundamental concepts that we must be able to fulfil in order to make up cryptography: encryption, decryption. At some point, we will also need key. This type of cryptography is called RSA cryptography.

Perhaps it will be easiest to explain with an example.

Let's say I want to send you an encrypted message.
To do this, you need to make a public key, which comprises two numbers, available to me.
This is where the prime numbers come in.
Let's say you choose two prime numbers: 11 and 17. In calculations, we call these numbers *p* and *q*. They must both be kept private, and make up your private key, which only you can access.
Multiply *p* and *q* to get 187. Let this be *N* — a public key number.
Your second public number is a smaller one, which you can choose. Let this be e.
There are some rules that restrict your choice, but let's say you pick 7. You can arrive at this choice by subtracting 1 from both p and q, and calculating pq. In this case, we will arrive at 160. Then, we must pick out a prime number that is not a factor of 160. 2 and 5 are not suitable, but 7 is suitable, so we can take the value of e as 7.
The encryption system has been set up, and N and e act as public keys, and can be shared with anyone. Conventionally, however, they are left in an online directory.
Now, let's say I wish to send the message '99'. Naturally, I have chosen a number for convenience, but any message can be represented numerically using binary, so the case will still stand.
99 is raised to the power of e, giving us (99x99x99x99x99x99x99). Divide this number by N (187) and calculate the remainder.

This can be perhaps more visually understood through the use of a clock, albeit a rather abstract one. If you construct a clock with 187 numbers, dividing the 99^e by N and taking the remainder would be the same as taking us around the clock 94 percent of the way. And, with that, we arrive at our answer, 176. The size of the clock will of course vary depending on N and your original prime numbers, but applying the relatively simple principle concept of remainders which mathematicians have overcomplicated by labelling 'modular arithmetic', we can figure out the encrypted message. We use remainders as they are simply much easier to work with than whole numbers, as processing large numbers to large exponents would be very challenging.



Then comes the final step of actually deciphering the message, 176.
We must introduce another number, d. We can calculate d by adding 1 to the product of (p-1)(q-1), giving us a result of, in this case, 161. Now divide this number by e, which we know to be 7, to get 23. We can arrive at this result from the Extended Euclidean Algorithm, and Fermat's Little Theorem can verify this statement. * You raise the message, 176 to the power of d, arriving at 176^23, divide by N and find the remainder. And, that should come out to 99. So, with only 187 and 7 ever being publicly available, we were able to transmit the message '99' without detection. Should anyone have intercepted the message '176', they would have no way of understanding what the message originally was.

Of course, using the number 187 would be far too easy for modern day computers to break down. In turn, perhaps my statement about '7' keeping your finances in tact was also hyperbolic. However, I think it becomes clear how, using 2 prime numbers that are hundreds -or even thousands of digits long- make finding the original prime numbers rather difficult. This of course also highlights the importance of the Greek's findings regarding the infinite amount of prime numbers, which prevent computers from being able to brute force the encryption by repeatedly using primes from a known list. Indeed, this means that many companies may use prime numbers that people don't even know exist to encrypt their messages.

However, just as prime numbers begun to get their time in the spotlight, the light may soon dwindle for them, for we are approaching the age of quantum computing. These machines- able to perform billions of calculations per second- will likely be able to crack RSA encryption with ease, making our current means of communication unsafe. Of course, researchers are already developing solutions to overcome this, but I nonetheless find it almost poetic that a concept discovered thousands of years ago finally found a random yet vital use, only to be shut down by rapidly evolving technology. Perhaps in the future, we will see prime numbers be used once again in some other way, but that may not be for thousands of years

However, I believe the most interesting question posed by this must be: what current, seemingly pointless mathematics, designed solely with the purpose of entertaining school kids or talented mathematicians may have a revolutionary use in the future? Happy numbers? The Collatz Conjecture? We often scrutinise the people 'wasting' their lives on this, but perhaps discoveries made about these problems will serve as the foundation for revolutionary mathematics in thousands of years.

*In truth, the proof is challenging for me to arrive at and fully comprehend. Please refer to (1) and (2) for further explanations.

References

(1) https://mathworld.wolfram.com/FermatsLittleTheorem.html

(2) http://www-math.ucdenver.edu/~wcherowi/courses/m5410/exeucalg.html

(3) https://youtu.be/JD72Ry60eP4

(4) https://www-users.york.ac.uk/~ss44/cyc/p/primeprf.htm

(5) https://www.redhat.com/en/blog/brief-history-cryptography

(6) https://www.abc.net.au/news/science/2018-01-20/how-prime-numbers-rsa-encryption-works/9338876